Data at NIST: A View from the Office of Data and Informatics

Robert Hanisch
Office of Data and Informatics
Material Measurement Laboratory
National Institute of Standards and Technology

NIST
National Institute of Standards and Technology
U.S. Department of Commerce

Data and NIST

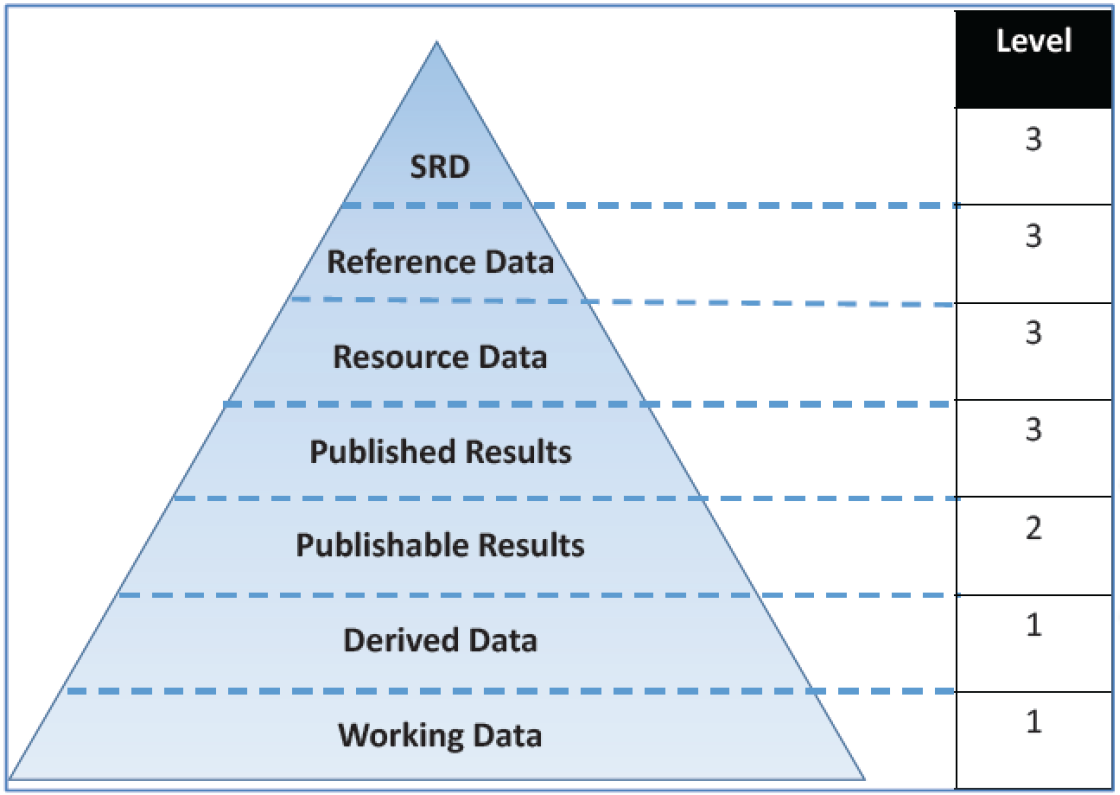
- NIST is a national and world resource for fundamental data
- Access should be easy and open
  - With regard to IP and privacy issues
- As our nation's standards organization…
  - NIST should be a leader in national and international standards efforts for data discovery and access
  - Discovery is fundamental
  - Discovery is enabled by metadata standards
- Key research at NIST should engage in data sharing strategies from the onset
- NIST should provide an infrastructure that makes data and information sharing as easy as possible

NIST Public Data Access Policy

- Establish NIST's commitment to providing public access to scientific research results
- Support governance of and best practices for managing peer-reviewed scholarly publications and digital scientific data across NIST
- Ensure effective access to and reliable preservation of NIST peer-reviewed scholarly publications and digital scientific data for use in research, development, education, and scientific discovery
- Enhance innovation and competitiveness by maximizing the potential to create new business opportunities

http://www.nist.gov/data/upload/NIST-Plan-for-Public-Access.pdf

NIST Public Data Access Policy



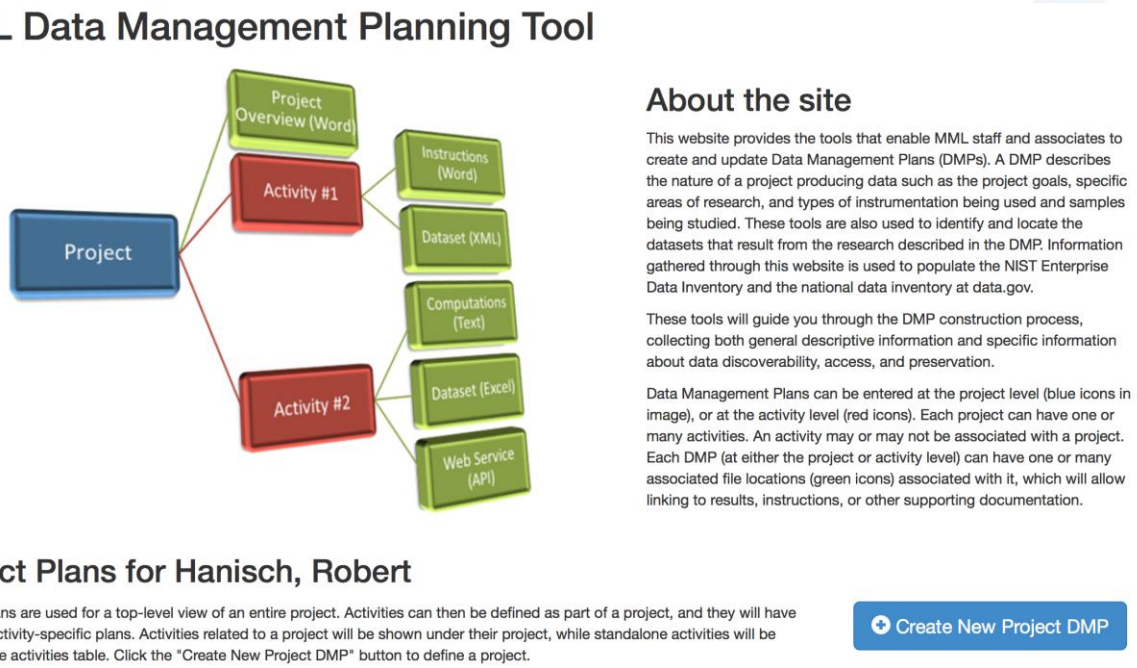
Metadata values in NIST EDI are made publicly available

Metadata goes to NIST EDI and PID minted

No requirements

Implementation

- Data management plans
- Enterprise Data Inventory
- data.gov



Data Management Plans

Listing Data Management Plans for Projects

Primary investigator	Title	Description	Category	Tags	Data category		
Linstrom, Peter	The NIST Chemistry Webbook	The NIST Chemistry WebBook provides users with easy access to chemical and physical property data for chemical species through the internet. The data provided in the site are from collections maintaine	Energy, Environment and Climate, Manufacturing, Safety, Security and Forensics	chemical data, thermochemical data, thermodynamic data, thermophysical data, enthalpy, entropy, heat capacity, heat of formation, chemical structure, ionization potential, thermochemistry, boiling point, vapor pressure, IR spectrum, mass spectrum, UV/Vis spectrum, retention index, InChI, InChIKey	Standard Reference Data (SRD)	Show	Edit
Scott, John Henry J.	Accelerated Discovery to Delivery -- SEM Data Formatting/Capture for DoD	As part of this project, I will assemble a collection of example SEM data files from different instrument manufacturers that will serve as exemplars of file formats and metadata. The data files will co	Advanced Materials	materials genome initiative, MGI, Army, Navy, Air Force, schema, metadata capture	Working Data	Show	Edit
Shen, Vincent K.	Molecular simulation of complex fluids	Simulation results related to research on complex fluids.	Advanced Materials, Biosciences and Health, Energy, Environment and Climate, Manufacturing	molecular dynamics, monte carlo, thermophysical fluid properties	Published Results	Show	Edit

Data Management Plans

Data Management Plan for "NIST/ARPA-E Database of Novel and Emerging Adsorbent Materials"

Summary of Activities

Title of Project: NIST/ARPA-E Database of Novel and Emerging Adsorbent Materials

Principal Investigator: Siderius, Daniel

Category: Advanced Materials, Energy, Environment and Climate, Manufacturing

Tags: adsorbate, adsorbent, adsorption, isotherm, metal organic Framework, porous Material, surface science

Organizational Code: 646.04 -- Chemical Informatics Research Group

Data Description (Data.gov): The NIST/ARPA-E Database of Novel and Emerging Adsorbent Materials (NIST SRD-205) is a free, web-based catalog of adsorbent materials and measured adsorption properties of numerous materials obtained from article entries from the scientific literature. Search fields for the database include adsorbent material, adsorbate gas, experimental conditions (pressure, temperature), and bibliographic information (author, title, journal), and results from queries are provided as a list of articles matching the search parameters. The database also contains adsorption isotherms digitized from the cataloged articles, which can be compared visually online in the web application or exported for offline analysis.

Process description (internal use): Data collection occurs through a sequence of steps: 1. PI compiles a list of articles from which to extract adsorption metadata and adsorption isotherms. Master list of articles abstracted or to-be-abstracted is maintained in an EndNote Library. 2. Students, interns, or NIST staff extract the correct adsorption metadata from articles, digitize adsorption isotherms, then input data into CSV flat files. 3. PI runs error checking software on CSV flat files, requests corrections. Steps 2 and 3 are repeated until CSV files pass error checks. 4. PI merges bibliographic information into metadata CSV file. 5. PI converts isotherm CSV files into JSON format. 6. PI uploads new dataset into MySQL database using administration panel (internal only) of web application. 7. Internal web server automatically send database updates to external web server once per week.

Release Date: 2014-10-28

Last updated: 2014-10-28

References: http://adsorbents.nist.gov, http://reaction.nist.gov/NISTOnly/adsorption_db/

Data Management Plans

Data Types and Classification

Data category (Preservation Level): Standard Reference Data (SRD) **All File Formats:** json, csv, xls

Data dictionary url: **Data dictionary type:** **Data standard:**

Data types description: Bibliographic data and article metadata is contained in CSV flat files. Isotherm data from articles are stored in JSON files will certain reserved attribute fields (DOI, source description [table or figure number], temperature, material name, gas name, pressure units, adsorption units, isotherm pressure/adsorption pairs). Other attribute fields may be added without breaking file functionality for SRD-205. No standard for this file format exists as of 02/23/2015.

Preservation

Backup method: File server, Other

Preservation description: Local backup - Dataset (raw and processed) and web application are stored on PI's desktop computer, which is automatically mirrored 3x/week to NAS fileserver http://h178112.nist.gov. Weekly backups (if data has changed) of processed SQL database on internal web server (http://reaction.nist.gov). SQL dumps are backed up 3x/week to NAS fileserver http://h178112.nist.gov. Web application code managed through private git repository (http://github.com/usnistgov/adsorption_db). Local version of web application code is backed up 3x/week to NAS fileserver http://h178112.nist.gov.

The following distributions are covered under this project:

Distribution Details	Download url	Media type	Version			
NIST/ARPA-E Database of Novel and Emerging Adsorbent Materials () The NIST/ARPA-E Database of Novel and Emerging Adsorbent Materials (NIST SRD-205) is a free, web-based catalog of adsorbent materials and measured adso...			1.0	Show	Edit	Remove

Describe New Distribution

Data Management Plans

Discoverablity and Access

Publisher: 646.04 -- Chemical Informatics Research Group

Relationship	Name	Organization		
Creator	Siderius, Daniel	Chemical Sciences Division	Edit	Remove
Contributor	Shen, Vincent K.	Chemical Sciences Division	Edit	Remove
Creator	van Zee, Roger D.	Chemical Sciences Division	Edit	Remove

Add NIST Staff Add External Staff

Homepage url: http://adsorbents.nist.gov

Language: en-US

Public access level: public

Rights: Dataset is Standard Reference Data and is covered by copyright under the Standard Reference Data Act. Data are freely available through interface website or via request to PI. Dataset (not database or application) may be licensed to Springer Publishing for use in Landolt-Bornstein Database.

License: TBD - depends on licensing agreement with Springer Publishing. The data are Standard Reference Data and are copyright by the U.S. Secretary of Commerce.

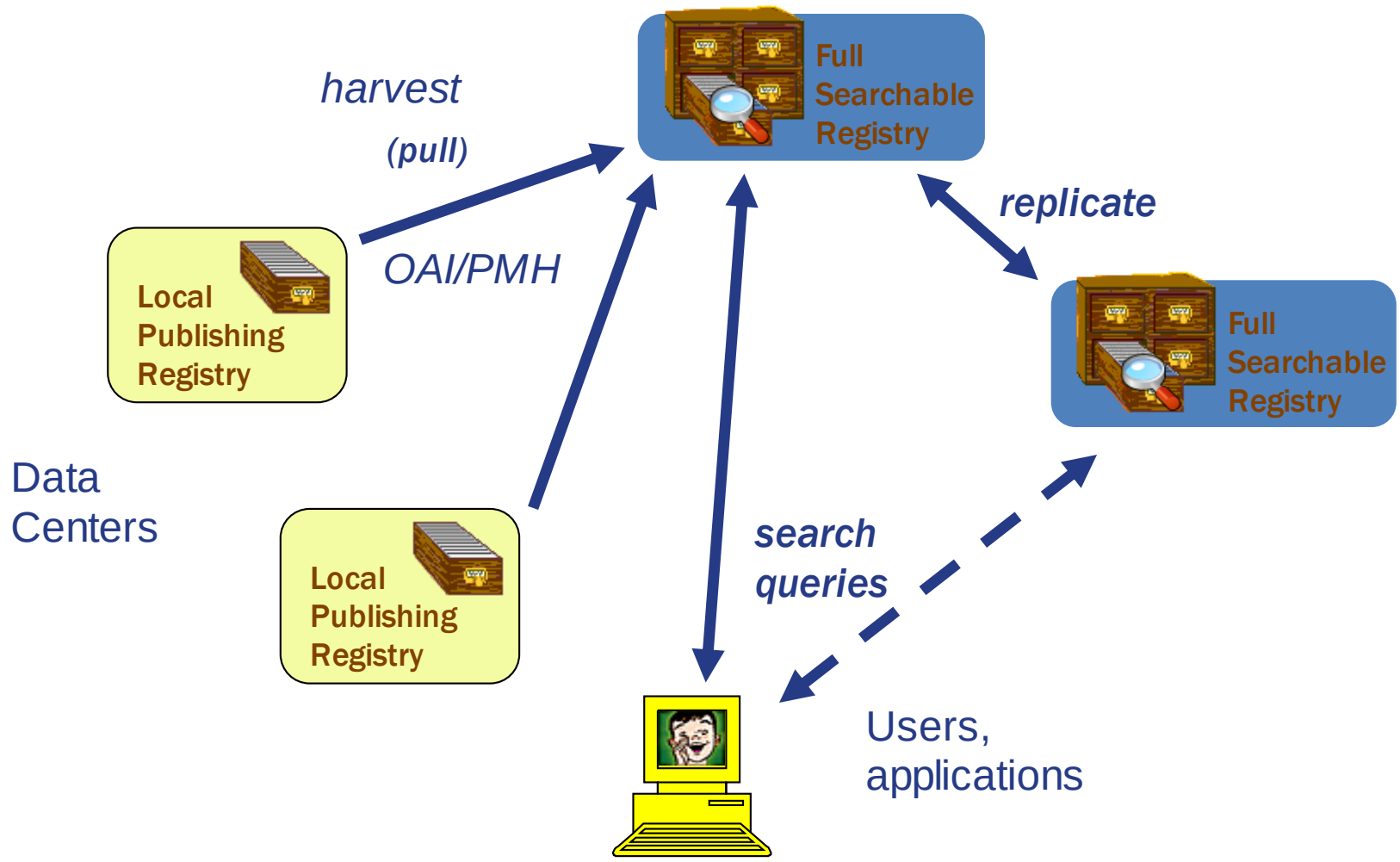
Discoverability access description: This SRD product exists as two parts: 1) the actual dataset (described in detail above) and 2) the PHP-based web application that is the interface to the MySQL database. The full web application, with administrative panel, resides on http://reaction.nist.gov, and is subject to version control using a private git repository on github.com. Currently, updates to the web application are manually pushed to the external server, http://adsorbents.nist.gov, by the PI. This could be improved by automating the update through a git pull that also deletes the administrative panel.

JSON Export to EDI, data.gov


```
    },
    {
        "title": "Cation substitution in thermochromic vanadium dioxide for smart windows",
        "identifier": "",
        "description": "This dataset includes infrared reflectances for thin film V_{1-x}M_{x}O_{2}, for M = Nb, Mo, W, Hf,
and x < 0.2, at temperatures of 5 \u00b0C to 85 \u00b0C, transition temperatures derived from the infrared reflectance
measurements, and x-ray diffraction spectra at 23 +- 3 \u00b0C. It also includes Matlab codes for analysis and presentation
of the data. The dataset supports a study of the depression of transition temperatures in lightly substituted vanadium
dioxide (VO_2) for smart energy-efficient building windows. While unsubstituted VO_2 undergoes a phase transition at 68
\u00b0C with concomitant changes in the infrared reflectances, the temperature of transition can be depressed by low level
cation substitution. ",
        "modified": "2015-02-19T11:32:35-05:00",
        "publisher": {
            "@type": "org:Organization",
            "name": "643.04  -- Functional Properties Group",
            "subOrganizationOf": {
                "@type": "org:Organization",
                "name": "National Institute of Standards and Technology",
                "subOrganizationOf": {
                    "@type": "org:Organization",
                    "name": "Department of Commerce",
                    "subOrganizationOf": {
                        "@type": "org:Organization",
                        "name": "U.S. Government"
                    }
                }
            }
        },
        "isPartOf": "",
        "accessLevel": "public",
        "keyword": [],
        "bureauCode": ["006:55"],
        "programCode": ["006:045"],
        "spatial": "NIST Gaithersburg",
        "theme": ["Advanced Materials", "Energy", "Environment and Climate"],
        "dataQuality": "true",
        "distribution": [
        ],
        "accrualPeriodicity": "",
        "landingPage": "",
        "language": ["en-US"]

    },
```

Federated Architecture



Office of Data and Informatics

SRD

- continue existing SRD distribution
- Quality Framework
- SRD Modes
- assess external need
- new product ideas
  - SRMDS
  - data streams
  - alternative delivery methods
- Open Data Initiative
- Open Govt Directive
- Data.gov

Research Data

- deal w/ data deluge
- provide advice to MML bench staff
- gather best practices
- interpret external rules & regulations
- reduce redundancy
- promote cooperation and coherent action
- manage changes in scholarly publishing
- coordinate with
  - WERB
  - Library
  - JResNIST

Lead/Liaison

- partner with ITL
- represent MML
  - NIST committees
  - NSTC & IWGs
  - NIH, NSF, DOE
  - other Fed Govt
  - Research Data Alliance (RDA)
- data standards
- champion proposals
  - budget initiatives
  - IMS
  - inter-agency, RDA

Data Science

The 4th paradigm?

- will it stand next to
  - theoretical
  - experimental
  - computational
- Cloud
- Statistical Learning
- Big Data
- Knowledge Discovery
- very fast growing
- *many* new jobs
- new degrees/depts

ODI in context

- NIST Material Measurement Laboratory
  - Materials Science and Engineering Division
  - Materials Measurement Services Division
  - Biosystems and Biomaterials Science Division
  - Biomolecular Measurement Division
  - Chemical Sciences Division
  - Applied Chemicals and Materials Division
  - Office of Reference Materials
  - Office of Data and Informatics

MML in context

- Associate Director for Laboratory Programs
  - Material Measurement Laboratory
  - Communications Technology Laboratory
  - Physical Measurement Laboratory
  - Engineering Laboratory
  - Information Technology Laboratory
  - Center for Nanoscale Science and Technology
  - NIST Center for Neutron Research

ODI Today: People

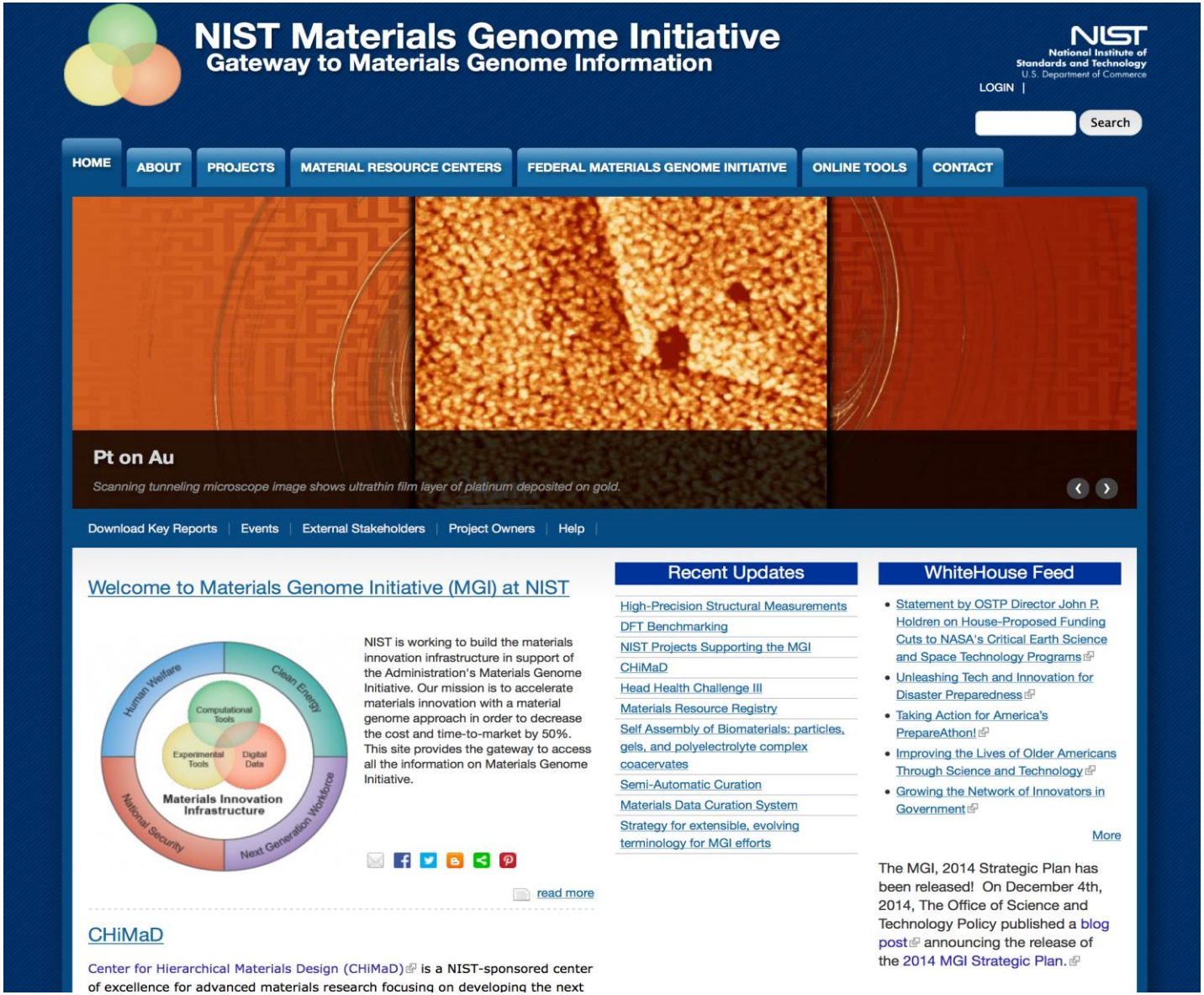
- Robert Hanisch, director
- Data Services Group Lead
  - Two web applications developers
  - SRD sales staff (3)
- Materials science detailee
- Chemistry detailee
- Biology detailee
- Data systems architect
- Data interoperability specialist
- Informatics/analytics consultant

- ODI Advisory Group
- SRD Advisory Group

ODI Today: Activities

- Data management plans, Enterprise Data Inventory (EDI)
- SRD modernization, audit, process
- MML Strategic Plan
- MML Data Working Group, Research Reproducibility user's group, electronic laboratory notebook evaluations, seminars, internal & external workshops
- NIST web site redesign, taxonomy
- NIST Library
- Materials Genome Initiative
  - materialsdata.nist.gov Dspace repository
  - Metadata standards
  - MGI community portal
- genomicsdata.nist.gov in development
- Software discovery, citation, re-use
- Nanomaterials registry, nanomaterials metadata standards, National Jet Fuels Program, NIST Center for Automotive Lightweighting, NIH collaboration
- InChI Trust membership
- Domain Repositories
  - Sustainability, interoperability
- DoC Data Working Group
  - DoC Data Council selection committee
- Research Data Alliance, National Data Services Consortium

Materials Genome Initiative



materialsdata.nist.gov

Material Measurement Laboratory DSpace Server

This is the NIST Material Measurement Laboratory DSpace server.

Use of this server is subject to terms of service

To get an account on this system (required for uploading), send a message to the administrator. Please include your requested username, e-mail address, and first and last name.

View our DSpace data repository.

Manage your credentials.

The National Institute of Standards and Technology (NIST) is an agency of the U.S. Department of Commerce.

Privacy Policy / Security Notice / Accessibility Statement / Disclaimer / Freedom of Information Act (FOIA) / Environmental Policy Statement / No Fear Act Policy / NIST Information Quality Standards / Scientific Integrity Summary

Date created: October 29, 2013 | Last updated: October 29, 2013 Contact: **Webmaster**

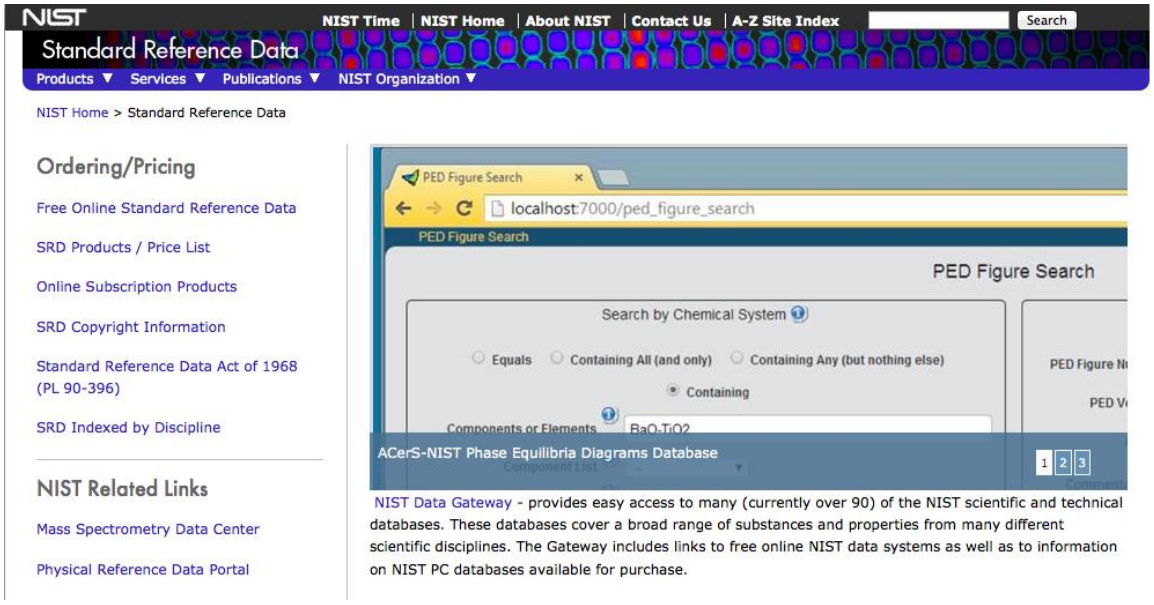
NIST
National Institute of Standards and Technology
U.S. Department of Commerce

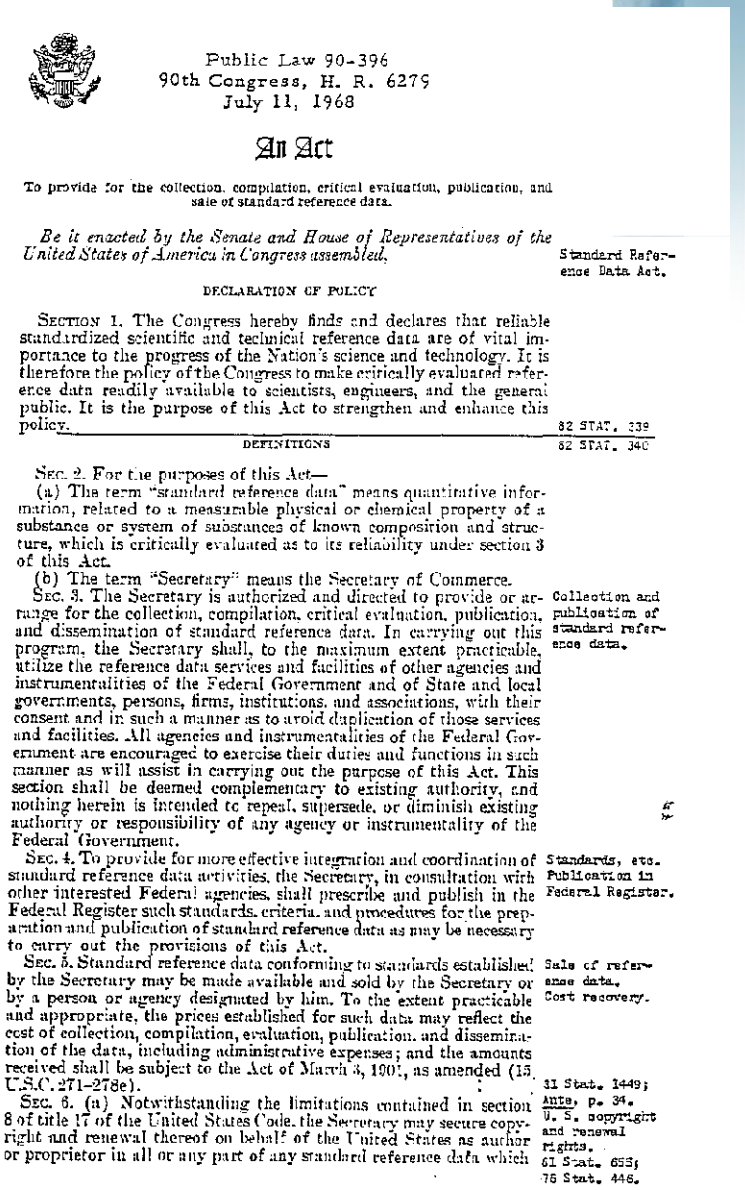
materialsdata.nist.gov

NIST
National Institute of Standards and Technology
U.S. Department of Commerce

Standard Reference Data

- SRD Act of 1968 authorized NIST to create Standard Reference Data
  - Copyright
  - Cost recovery
- 90 databases, most are free to use





Public Law 90-396
90th Congress, H. R. 6279
July 11, 1968

An Act

To provide for the collection, compilation, critical evaluation, publication, and sale of standard reference data.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

Standard Reference Data Act.

DECLARATION OF POLICY

SECTION 1. The Congress hereby finds and declares that reliable standardized scientific and technical reference data are of vital importance to the progress of the Nation's science and technology. It is therefore the policy of the Congress to make critically evaluated reference data readily available to scientists, engineers, and the general public. It is the purpose of this Act to strengthen and enhance this policy.

82 STAT. 339

DEFINITIONS

82 STAT. 340

SEC. 2. For the purposes of this Act—

(a) The term "standard reference data" means quantitative information, related to a measurable physical or chemical property of a substance or system of substances of known composition and structure, which is critically evaluated as to its reliability under section 3 of this Act.

(b) The term "Secretary" means the Secretary of Commerce.

Collection and publication of standard reference data.

SEC. 3. The Secretary is authorized and directed to provide or arrange for the collection, compilation, critical evaluation, publication, and dissemination of standard reference data. In carrying out this program, the Secretary shall, to the maximum extent practicable, utilize the reference data services and facilities of other agencies and instrumentalities of the Federal Government and of State and local governments, persons, firms, institutions, and associations, with their consent and in such a manner as to avoid duplication of those services and facilities. All agencies and instrumentalities of the Federal Government are encouraged to exercise their duties and functions in such manner as will assist in carrying out the purpose of this Act. This section shall be deemed complementary to existing authority, and nothing herein is intended to repeal, supersede, or diminish existing authority or responsibility of any agency or instrumentality of the Federal Government.

Standards, etc. Publication in Federal Register.

SEC. 4. To provide for more effective integration and coordination of standard reference data activities, the Secretary, in consultation with other interested Federal agencies, shall prescribe and publish in the Federal Register such standards, criteria, and procedures for the preparation and publication of standard reference data as may be necessary to carry out the provisions of this Act.

Sale of reference data. Cost recovery.

SEC. 5. Standard reference data conforming to standards established by the Secretary may be made available and sold by the Secretary or by a person or agency designated by him. To the extent practicable and appropriate, the prices established for such data may reflect the cost of collection, compilation, evaluation, publication, and dissemination of the data, including administrative expenses; and the amounts received shall be subject to the Act of March 3, 1901, as amended (15 U.S.C. 271–278e).

31 Stat. 1449; Ante, p. 34.

SEC. 6. (a) Notwithstanding the limitations contained in section 8 of title 17 of the United States Code, the Secretary may secure copyright and renewal thereof on behalf of the United States as author or proprietor in all or any part of any standard reference data which

U. S. copyright and renewal rights. 61 Stat. 653; 76 Stat. 446.

SRD Examples

NIST National Institute of Standards and Technology Physical Meas. Laboratory

XCOM

Element/Compound/Mixture Selection

In this database, it is possible to obtain photon cross section data for a single element, compound, or mixture (a combination of elements and compounds). Please fill out the following information:

Help

Identify material by:

- (•) Element
- ( ) Compound
- ( ) Mixture

Method of entering additional energies: (optional)

- (•) Enter additional energies by hand
- ( ) Additional energies from file (*Note: Your browser must be file-upload compatible*)

Submit Information | Reset

View Tables:

Please note that you can select only *ONE* table at a time of thermoelectric voltages of each type by temperature range, of the coefficients, or of the inverse coefficients.

Type	Temperature Range	Coefficients
B	Select Temperature Range	Select Coefficients Table
E	Select Temperature Range	Select Coefficients Table
J	Select Temperature Range	Select Coefficients Table
K	Select Temperature Range	Select Coefficients Table
N	Select Temperature Range	Select Coefficients Table
R	Select Temperature Range	Select Coefficients Table
S	Select Temperature Range	Select Coefficients Table
T	Select Temperature Range	Select Coefficients Table

Download Tables of Thermoelectric Voltages and Coefficients

View Thermocouple Types Definitions

View Corrections to Coefficients Tables

Go Back

Kinetics Database

NIST National Institute of Standards and Technology

Home
©NIST, 2013
Accessibility information

Kinetics Database Resources

Simple Reaction Search

Search Reaction Database

Search Bibliographic Database

Set Unit Preferences

Feedback

Submit an Article

Rate Our Products and

NIST Chemical Kinetics Database

Standard Reference Database 17, Version 7.0 (Web Version), Release 1.6.8
Data Version 2013.03

A compilation of kinetics data on gas-phase reactions

Notice: We are now accepting requests for abstracting kinetics data from journal articles and other references. Please use the "Submit an Article" link at the left if you find an article that has been missed in the database. You may request abstracting of a newer publication as well.

Reaction Database Quick Search Form

Enter the reactant(s) and/or product(s) in the fields below. Fields may be left blank.

[ ] + [ ] → [ ] + [ ] Submit Clear

If you would like more search options, try...
advanced reaction search form
bibliographic search form

ASD Data: Lines Levels — Information: Lines, Levels, Spectra, Ground States & Ionization Energies, Bibliography, Help

NIST National Institute of Standards and Technology Physical Meas. Laboratory

NIST Atomic Spectra Database Lines Form

Best viewed with the latest versions of Web browsers and JavaScript enabled

Spectrum: [ ] e.g., Fe I or Na;Mg; Al or mg i-iii
Lower Wavelength: [ ] or Upper Wavenumber (in cm⁻¹) [ ]
Upper Wavelength: [ ] or Lower Wavenumber (in cm⁻¹) [ ]
Units: nm

Reset Input | Retrieve Data

Dynamic Plots

Line Identification Plot: ( )
Saha-LTE Spectrum: ( )
Electron Temperature T_e(eV): [ ] Doppler-broadened spectrum
Electron Density N_e(cm⁻³): [ ] Ion Temperature T_i(eV): [ ] (if $T_i \neq T_e$)

Grotrian Diagram

Java subwindow size: ( ) 640 x 640 ( ) 800 x 640 (•) 1024 x 768 ( ) 1280 x 1024
[ ] Group by configurations [ ] Term multiplicity
[ ] Show only radiatively linked levels
Make Grotrian Diagram (requires Java2)

Output Options

Format output: HTML (formatted)
No JavaScript: [ ]
Energy Level Units: cm-1
Display output: in its entirety
Page size: 15
Output ordering: (•) Wavelength ( ) Multiplet

Additional Criteria

Lines: (•) All ( ) Only with transition probabilities ( ) Only with energy level classifications ( ) Only with observed wavelengths
Bibliographic Information: [✓] TP references, Line references
Wavelength Data: [✓] Observed [✓] Ritz [ ] Observed - Ritz (difference) [ ] Wavenumber (in cm⁻¹)

Data Infrastructure Investments

- NIST investing in infrastructure development to instigate changes in data management behaviors, assure compliance with OMB/OSTP policies on open data
- Four areas of focus
  - Internal open data processes and tools: data publication process, deployment of Data Management Plan and Enterprise Data Inventory tools in federated architecture
  - Data management infrastructure: cloud-based storage for working data, drag-and-drop user i/f and API, fast network link to cloud storage
  - Data dissemination and public access: NIST data portal, SRD interface re-design, data center infrastructure for SRD and other public-facing data
  - Effective utilization of commercial/public domain data tools: increase pace of review and potential adoption of data and collaboratory tools, move ~20 packages through NIST review processes (Skype, SpiderOak, OwnCloud, SciDrive, Socrata, Evernote, ELNs, etc.)

Research Data Alliance

Home

About

RDA
RESEARCH DATA ALLIANCE

Research Data Alliance

The Research Data Alliance (RDA) builds the social and technical bridges that enable open sharing of data Cross-border & cross-disciplinary challenges

The current global research data landscape is highly fragmented, by disciplines or by domains, from oceanography, life sciences and health, to agriculture, space and climate. When it comes to cross-disciplinary activities, the notions of "building blocks" of common data infrastructures and building specific "data bridges" are becoming accepted metaphors for approaching the data complexity and enable data sharing. The Research Data Alliance enables data to be shared across barriers through focused Working Groups and Interest Groups, formed of experts from around the world – from academia, industry and government. Participation in RDA is open to anyone who agrees to its guiding principles of openness, consensus, balance, harmonisation, community driven and non-profit approach. It was started in 2013 by a core group of interested agencies – the European Commission, the US National Science Foundation and National Institute of Standards and Technology, and the Australian Government's Department of Innovation. Other agencies, countries, companies, associations and institutes are due to join. RDA also has a broad, committed

RDA/CODATA Materials Data, Infrastructure & Interoperability IG

Status: Recognised & Endorsed

The development of advanced materials inherently rests on access to a distributed materials infrastructure and materials research data to fuel discovery and innovation.Given the complementary missions the RDA IG and the CODATA TG will work together under the following statement in support of the exchange of material data.

READ MORE 711 reads

http://rd-alliance.org/

CODATA

HOME | CODATA BLOG | EVENTS | NEWS | ABOUT CODATA | MEMBERS' AREA | CONTACT

CODATA

International Council for Science : Committee on Data for Science and Technology

Enter search terms... Search

Share:

ABOUT | EVENTS | MEMBERSHIP | COMMITTEES | TASK GROUPS | WORKING GROUPS | PUBLICATIONS | CONTACT | BLOG

The mission of CODATA is to strengthen international science for the benefit of society by promoting improved scientific and technical data management and use.

News & Articles

ISI CODATA International Training Workshop on Big Data

CODATA and the Indian Statistical Institute (ISI) - and other partners - will convene an International Training Workshop on Big Data. This will continue the emerging series of International Training Workshops which CODATA is seeking to establish with a variety of partners. Recent training workshops have been held in Beijing, China and in Nairobi, Kenya.

The International Training Workshop will take place at the Indian Statistical Institute in Bangalore, India, on 9-20 March 2015. Places for roughly twenty students will be available.

STATISTICAL INDIAN INSTITUTE UNITY IN DIVERSITY

Tweets Follow

Dawei Lin @iGenomics 8h

ICSU-Open access to scientific data and literature and the assessment of research by metrics icsu.org/general-assemb… #biocaddie #Force2015

Retweeted by CODATA

Expand

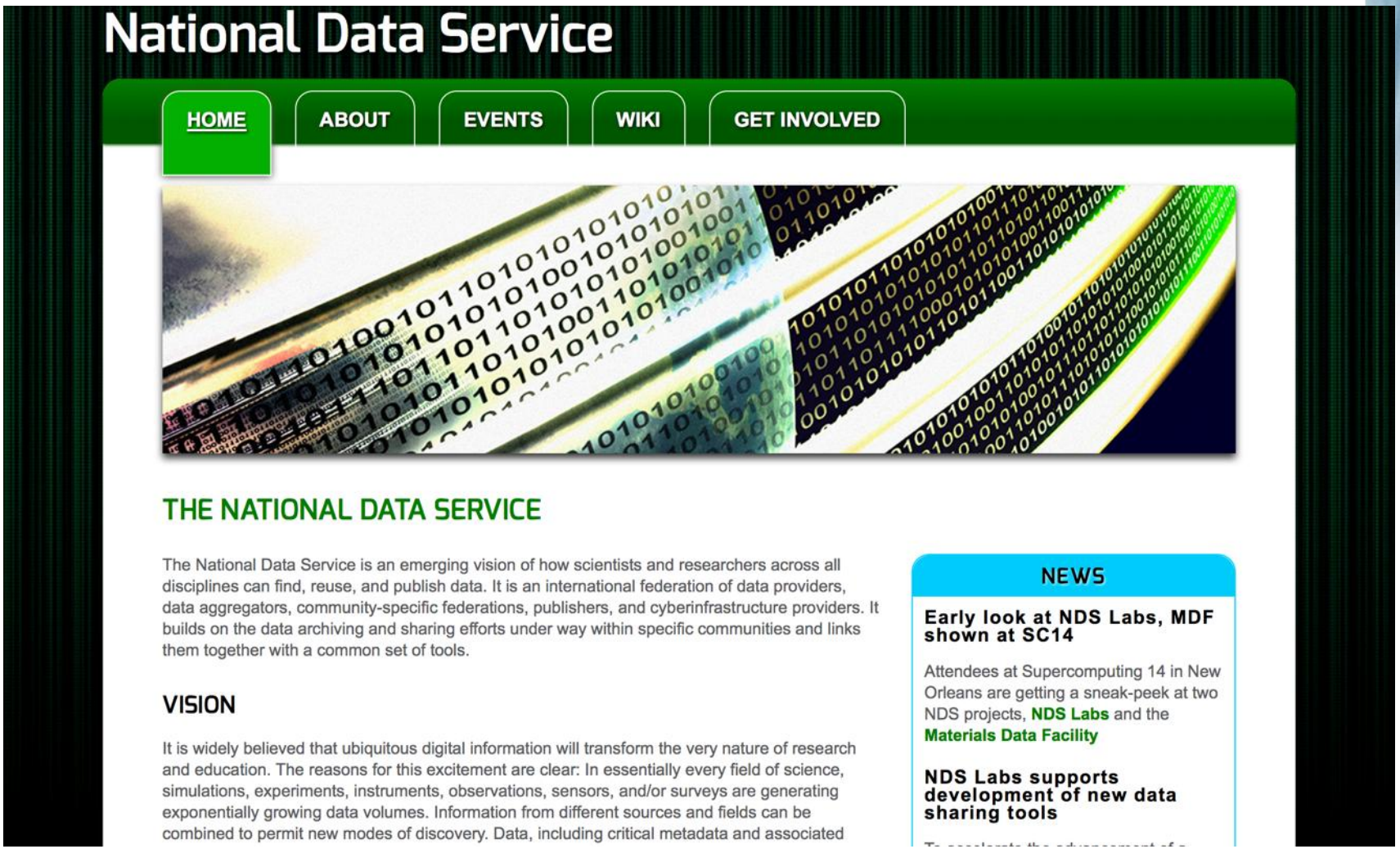
Dawei Lin @iGenomics 8h

@simonhodson99 Joint Declaration of Data Citation Principles force11.org/datacitation #biocaddie #Force2015

Retweeted by CODATA

http://www.codata.org/

National Data Service



http://www.nationaldataservice.org/

Hanisch, CENDI, 5/7/2015

International Data Week

- September 2016, DC area
- RDA Plenary, CODATA SciDataCon, ICSU World Data Service
- Sponsors being sought

NIST research data: 2014-2017

- Modernize the Standard Reference Data collection.
  - Revamp interfaces and build applications programming interfaces (APIs) in order to make SRDs much more usable.
  - Audit and rationalize SRD, special databases, etc.
- Implement the OMB/OSTP policies on providing public access to data.
  - Build a tool to facilitate creation of Data Management Plans by researchers and working to show the value of DMPs independent of the Administration directives.
  - Build a tool to allow researchers to deposit metadata for export to data.gov.
  - Develop production-level infrastructure and populate it with persistent identifiers and metadata for all publicly available NIST data.
- Establish a solution center for good data management practices.
  - Broker solutions to data management problems, advise on good practices and help staff set up the hardware/software infrastructure needed to support their data management challenges.
- Become a resource for data analytics and informatics solutions.
  - Provide consultation from an informatics generalist who can help researchers in a variety of situations to find the most appropriate tools for understanding the patterns and characteristics of complex/large data sets.
- Support the Materials Genome Initiative
  - National initiative to promote new materials—discovery through deployment—by linking data, models, and experiment.

NIST research data: ~10 year horizon

- Expand the Standard Reference Data collection.
  - Identify through internal and external inputs where new SRD are needed.
  - Prioritize, scope, and find resources for development work
- Establish NIST as an exemplar federal agency in data management.
  - Implement and share best practices for preservation, curation, discovery, re-use, and interoperability
  - Facilitate community-based development of metadata standards & data models
  - Participate in leadership of national and international data federation activities
    - Research Data Alliance, National Data Services Consortium, CODATA and World Data System
  - Contribute to solving the challenge of long-term sustainability of data repositories
  - Share NIST-developed technologies to assist other agencies in improving data access and data services
  - Collaborate with federal and non-federal organizations in developing and deploying common solutions
  - Establish a data-aware, data-savvy culture at NIST
    - Improve efficiency of experimentation and simulation
    - Improve reliability and reproducibility of research results
    - Increase value of NIST to the research and industrial communities

Summary

- Goal of ODI is to help all MML scientists
  - Data management plans
  - Best practices
  - Improved efficiency: do science, not bookkeeping
  - Data publication and citation
- But… MML data are incredibly diverse
  - New metadata standards
  - Must address interoperabiity at the appropriate granularity
- Can utilize many elements of the Virtual Observatory infrastructure and architecture to improve data management capabilities in MML and NIST
  - Federation rather than centralization

Some things to think/worry about

- Quality metadata is key for discovery, interoperability, re-use
  - Reproducibility
  - Integrity of the scientific process
  - Metadata curation is non-trivial, can be costly
- Address interoperability at the proper scale
  - Too wide: very expensive, difficult/impossible to reach consensus across disciplines; what is the scientific motivation?
  - Too narrow: Scientific stovepipes, missed opportunities for discovery at the intersections of complementary data collections

Some things to think/worry about

- Standards for metadata, data access protocols, etc., require community participation to assure take-up
  - Major research organizations
  - Professional societies (national, international)
  - Recognized standards organizations
  - RDA, CODATA, NDS, EUDAT, etc.

Some things to think/worry about

- Little national commitment to sustaining infrastructure for open data
  - Domain repositories often must (re)compete for basic resources, rely on complex business models
  - Federal funding agencies require Data Management Plans, but provide no common infrastructure and no consistent review process
  - Commercial academic publishers poised to take on data preservation roles; open data could move behind pay-walls

http://tinyurl.com/domainrepositories25